GARC: A New Associative Classification Approach

I. Bouzouita, S. Elloumi, and S. Ben Yahia

Faculty of Sciences of Tunis,
Computer Science Department, 1060 Tunis, Tunisia
{samir.elloumi, sadok.benyahia}@fst.rnu.tn

Abstract. Many studies in data mining have proposed a new classification approach called *associative classification*. According to several reports associative classification achieves higher classification accuracy than do traditional classification approaches. However, the associative classification suffers from a major drawback: it is based on the use of a very large number of classification rules; and consequently takes efforts to select the best ones in order to construct the classifier. To overcome such drawback, we propose a new associative classification method called GARC that exploits a generic basis of association rules in order to reduce the number of association rules without jeopardizing the classification accuracy. Moreover, GARC proposes a new selection criterion called *score*, allowing to ameliorate the selection of the best rules during classification. Carried out experiments on 12 benchmark data sets indicate that GARC is highly competitive in terms of accuracy in comparison with popular associative classification methods.

Keywords: Associative Classification, Generic Basis, Classification Rules, Generic association rules, Classifier.

1 Introduction

In the last decade, a new approach called *associative classification* (AC) was proposed to integrate association rule mining and classification in order to handle large databases. Given a training data set, the task of an associative classification algorithm is to discover the classification rules which satisfy the user specified constraints denoted respectively by minimum support (*minsup*) and minimum confidence (*minconf*) thresholds. The classifier is built by choosing a subset of the generated classification rules that could be of use to classify new objects or instances. Many studies have shown that AC often achieves better accuracy than do traditional classification techniques [1,2]. In fact, it could discover interesting rules omitted by well known approaches such as C4.5 [3]. However, the main drawback of this approach is that the number of generated associative classification rules could be large and takes efforts to retrieve, prune, sort and select high quality rules among them. To overcome this problem, we propose a new approach called GARC which uses generic bases of association rules. The

A Min Tjoa and J. Trujillo (Eds.): DaWaK 2006, LNCS 4081, pp. 554–565, 2006.

main originality of GARC is that it extracts the generic classification rules directly from a generic basis of association rules, in order to retain a small set of rules with higher quality and lower redundancy in comparison with current AC approaches. Moreover, a new score is defined by the GARC approach to find an effective rule selection during the class label prediction of a new instance, in the sake of reducing the error rate. This tackled issue is quite challenging, since the goal is to use generic rules while maintaining a high classifier accuracy.

The remainder of the paper is organized as follows. Section 2 briefly reports basic concepts of associative classification and scrutinizes related pioneering works. Generic bases of association rules are surveyed in section 3. Section 4 presents our proposed approach, where details about classification rules discovery, building classifier and prediction of test instances are discussed. Experimental results and comparisons are given in section 5. Finally, section 6 concludes this paper and points out future perspectives.

2 Associative Classification

An association rule is a relation between itemsets having the following form: $R : X \Rightarrow Y - X$, where X and Y are frequent itemsets for a minimal support *minsup*, and $X \subset Y$. Itemsets X and $(Y-X)$ are called, respectively, *premise* and *conclusion* of the rule R. An association rule is valid whenever its strength metric, confidence(R)=$\frac{support(Y)}{support(X)}$, is greater than or equal to the minimal threshold of confidence *minconf*.

An associative classification rule (ACR) is a special case of an association rule. In fact, an ACR conclusion part is reduced to a single item referring a class attribute. For example, in an ACR such as $X \Rightarrow c_i$, c_i must be a class attribute.

2.1 Basic Notions

Let us define the classification problem in an association rule task. Let D be the training set with n attributes (columns) A_1, .., A_n and $|D|$ rows. Let C be the list of class attributes.

Definition 1. *An object or instance in D can be described as a combination of attribute names and values a_i and an attribute class denoted by c_i [4].*

Definition 2. *An item is described as an attribute name and a value a_i [4].*

Definition 3. *An itemset can be described as a set of items contained in an object.*

A classifier is a set of rules of the form $A_1, A_2, ..., A_n \Rightarrow c_i$ where A_i is an attribute and c_i is a class attribute. The classifier should be able to predict, as accurately as possible, the class of an unseen object belonging to the test data set. In fact, it should maximise the equality between the predicted class and the hidden actual class.

The AC achieves higher classification accuracy than do traditional classification approaches [1,2]. The classification model is a set of rules easily understandable by humans and that can be edited [1,2].

2.2 Related Work

One of the first algorithms to use association rule approach for classification was CBA [4]. CBA, firstly, generates all the association rules with certain support and confidence thresholds as candidate rules by implementing the Apriori algorithm [5]. Then, it selects a small set from them by evaluating all the generated rules against the training data set. When predicting the class attribute for an example, the highest confidence rule, whose the body is satisfied by the example, is chosen for prediction.

CMAR [6] generates rules in a similar way as CBA with the exception that CMAR introduces a CR-tree structure to handle the set of generated rules and uses a set of them to make a prediction using a weighted $\chi 2$ metric [6]. The latter metric evaluates the correlation between the rules.

ARC-AC and ARC-BC have been introduced in [7,8] in the aim of text categorization. They generate rules similar to the Apriori algorithm and rank them in the same way as do CBA rules ranking method. ARC-AC and ARC-BC calculate the average confidence of each set of rules grouped by class attribute in the conclusion part and select the class attribute of the group with the highest confidence average.

The CPAR [2] algorithm adopts FOIL [9] strategy in generating rules from data sets. It seeks for the best rule itemset that brings the highest gain value among the available ones in data set. Once the itemset is identified, the examples satisfying it will be deleted until all the examples of the data set are covered. The searching process for the best rule itemset is a time consuming process, since the gain for every possible item needs to be calculated in order to determine the best item gain. During rule generation step, CPAR derives not only the best itemset but all close similar ones. It has been claimed that CPAR improves the classification accuracy whenever compared to popular associative methods like CBA and CMAR [2].

A new AC approach called Harmony was proposed in [10]. Harmony uses an instance-centric rule generation to discover the highest confidence discovering rules. Then, Harmony groups the set of rules into k groups according to their rule conclusions, where k is the total number of distinct class attributes in the training set. Within the same group of rules, Harmony sorts the rules in the same order as do CBA. To classify a new test instance, Harmony computes a score for each group of rules and assign the class attribute with the highest score or a set of class attributes if the underlying classification is a multi-class problem. It has been claimed that Harmony improves the efficiency of the rule generation process and the classification accuracy if compared to CPAR [2].

The main problem with AC approaches is that they generate an overwhelming number of rules during the learning stage. In order to overcome this drawback, our proposed approach tries to gouge this fact by the use of generic bases of association rules in the classification framework. In the following, we begin by recall some key notions about the Formal Concept Analysis (FCA), a mathematical tool necessary for the derivation of generic bases of association rules.

3 Generic Bases of Association Rules

The problem of the relevance and usefulness of extracted association rules is of primary importance. Indeed, in most real life databases, thousands and even millions of highly confident rules are generated among which many are redundant. In the following, we are interested in the lossless information reduction of association rules, which is based on the extraction of a generic subset of all association rules, called *generic basis* from which the remaining (redundant) association rules may be derived. In the following, we will present the generic basis of Bastide *et al.* [11,12] and $\mathcal{IGB}$ [13] after a brief description of FCA mathematical background necessary for the derivation of generic bases of association rules.

3.1 Mathematical Background

Interested reader for key results from the Galois lattice-based paradigm in FCA is referred to [14].

Formal context: A formal context is a triplet $\mathcal{K} = (\mathcal{O}, \mathcal{I}, \mathcal{R})$, where $\mathcal{O}$ represents a finite set of transactions, $\mathcal{I}$ is a finite set of items and $\mathcal{R}$ is a binary (incidence) relation (*i.e.*, $\mathcal{R} \subseteq \mathcal{O} \times \mathcal{I}$). Each couple $(o, i) \in \mathcal{R}$ expresses that the transaction $o \in \mathcal{O}$ contains the item $i \in \mathcal{I}$.

Frequent closed itemset: An itemset $I \subseteq \mathcal{I}$ is said to be *closed* if $\omega(I) = I^{(1)}$ [15]. I is said to be *frequent* if its *relative support*, Support$(I) = \frac{|\psi(I)|}{|\mathcal{O}|}$, exceeds a user-defined minimum threshold, denoted *minsup*.

Minimal generator [12]**:** An itemset $g \subseteq \mathcal{I}$ is said to be *minimal generator* of a closed itemset f, if and only if $\omega(g) = f$ and does not exist $g_1 \subseteq g$ such that $\omega(g_1) = f$. The set $\mathcal{G}_f$ of the minimal generators of f is: $\mathcal{G}_f = \{g \subseteq \mathcal{I} \mid \omega(g) = f \wedge \nexists\, g_1 \subset g$ such as $\omega(g_1) = f\}$.

3.2 The Generic Basis for Exact Association Rules ($\mathcal{GBE}$) and the Informative Basis for Approximate Association Rules ($\mathcal{GBA}$)

Bastide *et al.* considered the following rule-redundancy definition [12]:

Definition 4. *Let $\mathcal{AR}$ be a set of association rules derived from an extraction context $\mathcal{K}$ and c be a confidence value. A rule R: $X \overset{c}{\Rightarrow} Y \in \mathcal{AR}$ is redundant in comparison with R_1: $X_1 \overset{c}{\Rightarrow} Y_1$ if R fulfills the following constraints:*

1. *Support(R) = Support(R_1) and Confidence(R) = Confidence(R_1)= c;*
2. *$X_1 \subseteq X \wedge Y \subset Y_1$.*

The *generic basis for exact association rules* is defined as follows:

Definition 5. *Let $\mathcal{FCI}_{\mathcal{K}}$ be the set of frequent closed itemsets extracted from the extraction context $\mathcal{K}$. For each frequent closed itemset $f \in \mathcal{FCI}_{\mathcal{K}}$, let $\mathcal{G}_f$ be the set of its minimal generators. The generic basis of exact association rules*

[1] The closure operator is indicated by ω.

$\mathcal{GBE}$ is given by: $\mathcal{GBE} = \{R\text{: } g \Rightarrow (f - g) \mid f \in \mathcal{FCI}_{\mathcal{K}}$ and $g \in \mathcal{G}_f$ and $g \neq f^{(2)}\}$.

Bastide *et al.* also characterized the informative basis for approximate association rules, defined as follows [12]:

Definition 6. *Let $\mathcal{FCI}_{\mathcal{K}}$ be the set of frequent closed itemsets extracted from the extraction context $\mathcal{K}$. The $\mathcal{GBA}$ basis is defined as follows [12]:*
$\mathcal{GBA} = \{R \mid R\text{: } g \Rightarrow (f_1 - g) \mid f, f_1 \in \mathcal{FCI}_{\mathcal{K}}$ and $\omega(g) = f$ and $f \preceq f_1$ and Confidence(R) $\geq$ minconf $\}$.

The pair ($\mathcal{GBE}$, $\mathcal{GBA}$) is informative, sound and lossless [12,16] and rules belonging to this pair are referred as *informative association rules.*

3.3 Informative Generic Basis ($\mathcal{IGB}$)

The $\mathcal{IGB}$ basis is defined as follows:

Definition 7. *Let $\mathcal{FCI}_{\mathcal{K}}$ be the set of frequent closed itemsets and $\mathcal{G}_f$ be the set of minimal generators of all the frequent itemsets included or equal to a closed frequent itemset f. The $\mathcal{IGB}$ basis is defined as follows [13]:*

$\mathcal{IGB} = \{R\text{: } g_s \Rightarrow (f_1 - g_s) \mid f, f_1 \in \mathcal{FCI}_{\mathcal{K}}$ and $(f - g_s) \neq \emptyset$ and $g_s \in \mathcal{G}_f \wedge f_1 \preceq f \wedge$ confidence(R) $\geq$ minconf $\wedge \nexists\, g' \subset g_s$ such that confidence($g' \Rightarrow f_1$-g')$\geq$ minconf$\}$.

$\mathcal{IGB}$ basis [13] presents the following characteristics:

1. **Conveying maximum of useful knowledge**: Association rules of the $\mathcal{IGB}$ basis convey the maximum of useful knowledge. Indeed, a generic association rule of $\mathcal{IGB}$ is based on a frequent closed itemset and has the minimal premise since the latter is represented by one of the smallest frequent minimal generators satisfying *minconf* threshold. It was shown that this type of association rules conveys the maximum of useful knowledge [17];
2. **Information lossless:** It was pointed out that the $\mathcal{IGB}$ basis is extracted without information loss [13];
3. **Compactness**:the $\mathcal{IGB}$ basis is more compact than other informative generic basis [13], *e.g.*, the pair ($\mathcal{GBE}$, $\mathcal{GBA}$).

4 GARC: A New Associative Classification Approach

In this section, we propose a new AC method GARC[3] that extracts the generic classification rules directly from a generic basis of association rules in order to overcome the drawback of the current AC approaches, i.e., the generation of a large number of associative classification rules. In the following, we will present and explain in details the GARC approach.

[2] The condition $g \neq f$ ensures discarding non-informative rules of the form $g \Rightarrow \emptyset$.
[3] The acronym GARC stands for: Generic Association Rules based Classifier.

4.1 Rule Generation

In this step, GARC extracts the generic basis of association rules. Once obtained, generic rules are filtered out to retain only rules whose conclusions include a class attribute. Then, by applying the decomposition axiom, we obtain new rules of the form $A_1, A_2, ..., A_n \Rightarrow c_i$. Even though, the obtained rules are redundant, their generation is mandatory to guarantee a maximal cover of the necessary rules.

The $\mathcal{IGB}$ basis is composed of rules with a small premise which is an advantage for the classification framework when the rules imply the same class. For example, let us consider two rules R_1: `A B C D` $\Rightarrow$`cl1` and R_2: `B C` $\Rightarrow$`cl1`. R_1 and R_2 have the same attribute conclusion. R_2 is considered to be more interesting than R_1, since it is needless to satisfy the properties `A D` to choose the class `cl1`. Hence, R_2 implies less constraints and can match more objects of a given population than R_1.

Let us consider a new object O_x: `B C D`. If we have in the classifier just the rule R_1, we cannot classify O_x because the attribute A does not permit the matching. However, the rule R_2, which has a smaller premise than R_1, can classify O_x. This example shows the importance of the generic rules and, especially, the use of the $\mathcal{IGB}$ basis to extract the generic classification rules. In fact, such set of rules is smaller than the number of all the classification rules and their use is benefical for classifying new objects.

4.2 Classifier Builder

Once the generic classification rules obtained, a total order on rules is set as follows. Given two rules R_1 and R_2, R_1 is said to precede R_2, denoted $R_1 > R_2$ if the followed condition is fulfilled:

- $confidence(R_1) > confidence(R_2)$ or
- $confidence(R_1) = confidence(R_2)$ and $support(R_1) > support(R_2)$ or
- $confidence(R_1) = confidence(R_2)$ and $support(R_1) = support(R_2)$ and R_1 is generated before R_2.

The data set coverage is similar to that in CBA. In fact, a data object of the training set is removed after it is covered by a selected generic rule.

The major difference with current AC approaches [4,6,7,8,10] is that we use generic ACR directly deduced from generic bases of association rules to learn the classifier as shown by algorithm 1.

4.3 New Instance Classification

After a set of rules is selected for classification, GARC is ready to classify new objects. Some methods such as those described in [4,7,8,10] are based on the support-confidence order to classify a new object. However, the confidence measure selection could be misleading, since it may identify a rule `A` $\Rightarrow$ `B` as an interesting one even though, the occurrence of `A` does not imply the occurrence of `B` [18]. In fact, the confidence can be deceiving since it is only an estimate of

```
Data: 𝒟: Training data, 𝒢ℛ: a set of generic classification rules
Results: 𝒞: Classifier
Begin
    𝒢ℛ=sort(𝒢ℛ) in a descending order;
    Foreach rule r ∈ 𝒢ℛ do
        Foreach object d ∈ 𝒟 do
            If d matches r.premise then
                remove d from 𝒟 and mark r if it correctly classifies d;
            If r is marked then
                insert r at the end of 𝒞;
                select a default class;
    add the default class at the end of the classifier;
    return Classifier 𝒞 ;
End
```

Algorithm 1: Garc: *selected generic rules based on database coverage*

the conditional probability of itemset B given an itemset A and does not measure the actual strength of the implication between A and B. Let us consider the example shown in Table 1 which shows the association between an item A and a class attribute B. A and $\overline{\mathtt{A}}$ represent respectively the presence and absence of item A, B represents a class attribute and $\overline{\mathtt{B}}$ the complement of B. We consider the associative classification $\mathtt{A} \Rightarrow \mathtt{B}$. The confidence of this rule is given by *confidence*$(\mathtt{A} \Rightarrow \mathtt{B})=\frac{support(\mathtt{AB})}{support(\mathtt{A})} = \frac{201}{250} = 80.4\%$. Hence, this rule has high confidence. Now, let us calculate the correlation between A and B by using the lift metric [18]. $lift(\mathtt{A} \Rightarrow \mathtt{B})=\frac{support(\mathtt{AB})}{support(\mathtt{A})*support(\mathtt{B})} = \frac{0.201}{0.250*0.900} = 0.893$. The fact that this quantity is less than 1 indicates negative correlation between A and B.

Table 1. Example

	B	$\overline{\mathtt{B}}$	Total
A	201	49	250
$\overline{\mathtt{A}}$	699	51	750
Total	900	100	1000

To avoid the lacuna of using only confidence metric, we define a new lift based score formula as follows:

$$Score = \frac{1}{|\mathtt{Premise}|} * lift^{\frac{|\mathtt{Premise}|}{numberofitems}} = \frac{1}{|\mathtt{Premise}|} * \left(\frac{support(\mathtt{Rule})}{support(\mathtt{Premise}) * support(\mathtt{Conclusion})}\right)^{\frac{|\mathtt{Premise}|}{numberofitems}}$$

The introduced score includes the lift metric. In fact, the lift finds interesting relationships between A and B. It computes the correlation between the occurrence of A and B by measuring the real strength of the implication between them which is interesting for the classification framework. Moreover, the lift is divided by the cardinality of the rule premise part in order to give a preference to rules with small premises. Thus, Garc collects the subset of rules matching the new object attributes from the classifier. Trivially, if all the rules matching it have the same class, GARC just assigns that class to the new object. If the rules do not imply the same class attribute, the score firing is computed for each rule. The rule with the highest score value is selected to classify the new object.

5 Experiments

We have conducted experiments to evaluate the accuracy of our proposed approach GARC, developed in C++, and compared it to the well known classifiers CBA, ID3, C4.5 and Harmony. Experiments were conducted using 12 data sets taken from UCI Machine Learning Repository[(4)]. The chosen data sets were discretized using the LUCS-KDD [(5)] software.

The features of these data sets are summarized in Table 2. All the experiments were performed on a 2.4 GHz Pentium IV PC under Redhat Linux 7.2.

Table 2. Data set description

Data set	# attributes	# transactions	# classes
Monks1	6	124	2
Monks2	6	169	2
Monks3	6	122	2
Spect	23	80	2
Pima	38	768	2
TicTacToe	29	958	2
Zoo	42	101	7
Iris	19	150	3
Wine	68	178	3
Glass	48	214	7
Flare	39	1389	9
Pageblocks	46	5473	5

Classification accuracy can be used to evaluate the performance of classification methods. It is the percentage of correctly classified examples in the test set and can be measured by splitting the data sets into a training set and a test set.

During experiments, we have used available test sets for data sets Monks1, Monks2 and Monks3 and we applied the 10 cross-validation for the rest of data sets, in which a data set is divided into 10 subsets; each subset is in turn used as testing data while the remaining data is used as the training data set; then the average accuracy across all 10 trials is reported.

The parameters are set as the following. In the rule generation algorithm, *minsup* is set to 10% and *minconf* to 80%. In order to extract generic association

[4] *Available at* http://www.ics.uci.edu/∼mlearn/MLRepository.html

[5] *Available at* http://www.csc.liv.ac.uk/∼frans/KDD/Software/LUCS-KDD-DN/lucs-kdd DN.html

rules, we used the Prince algorithm [19] to generate both the pair ($\mathcal{GBE}$, $\mathcal{GBA}$) and $\mathcal{IGB}$ bases.

To evaluate C4.5 and ID3, we used the Weka(6) software and the Harmony prototype was kindly provided by its authors. We have implemented the CBA algorithm in C++ under Linux.

In the following, we will compare the effectiveness of the use of generic bases of the pair ($\mathcal{GBE}$, $\mathcal{GBA}$) and $\mathcal{IGB}$ for the classification framework. For this, we conducted experiments with reference to accuracy in order to compare the classifiers Garc$_B$ and Garc$_I$ issued respectively from the generic bases of the pair ($\mathcal{GBE}$, $\mathcal{GBA}$) and $\mathcal{IGB}$ without using the score firing.

Moreover, to show the impact of the score firing on the quality of the produced classifiers, we report the accuracy results of Garcs$_B$ and Garc deduced respectively from the generic bases of the pair ($\mathcal{GBE}$, $\mathcal{GBA}$) and $\mathcal{IGB}$ using the score firing.

5.1 The Score Firing Impact

Table 3 represents a comparison between the classifiers deduced from the generic bases of the pair ($\mathcal{GBE}$, $\mathcal{GBA}$) and $\mathcal{IGB}$ when using or not the score firing.

Table 3. Accuracy comparison of Garc$_B$, Garc$_I$, Garcs$_B$ and Garc algorithms for *minsup*=10% and *minconf*=80%

	Without using the score		Using the score	
Data set	Garc$_B$	Garc$_I$	Garcs$_B$	Garc
Monks1	92.0	92.0	92.0	92.0
Monks2	56.0	56.0	56.0	56.0
Monks3	96.3	96.3	96.3	96.3
Spect	67.0	68.9	67.0	68.9
Pima	73.0	73.0	73.0	73.0
TicTacToe	65.0	67.4	65.0	65.0
Zoo	89.0	89.0	89.0	90.0
Iris	95.0	94.7	95.6	95.4
Wine	89.2	89.4	90.0	89.8
Glass	58.0	59.3	58.0	64.0
Flare	85.0	85.0	85.0	85.0
Pageblocks	92.0	89.8	92.0	89.8
Average accuracy	79.7	80.0	79.9	**80.4**

Table 3 points out that the use of the score firing increases the accuracy performance for the classifiers deduced from the pair ($\mathcal{GBE}$, $\mathcal{GBA}$). In fact, Garcs$_B$ has a better average accuracy than Garc$_B$. Moreover, for the classifiers deduced from $\mathcal{IGB}$, the use of the score firing ameliorates the accuracy for four data sets. In fact, Garc outperforms Garc$_I$ on Zoo, Iris, Wine and Glass data sets.

[6] *Available at* http://www.cs.waikato.ac.nz/ml/Weka

Thus, the best average accuracy, highlighted in bold print, is given by GARC. Furthermore, as shown in Table 4, the number of rules generated by GARC is less than that generated by the approaches deduced from the pair ($\mathcal{GBE}$, $\mathcal{GBA}$), *i.e.*, GARC$_B$ and GARCS$_B$. In the following, we put the focus on comparing GARC accuracy versus that of the well known classifiers ID3, C4.5, CBA and Harmony.

Table 4. Number of associative classification rules for *minsup*=10% and *minconf*=80%

Data set	# generic ACR deduced from $\mathcal{IGB}$	# generic ACR deduced from ($\mathcal{GBE}$, $\mathcal{GBA}$)
Monks1	12	12
Monks2	4	4
Monks3	20	20
Pima	20	20
TicTacToe	15	15
Zoo	832	1071
Iris	22	24
Wine	329	471
Glass	31	36
Flare	237	561
Pageblocks	128	128

5.2 Generic Classification Rules Impact

Table 5 represents the accuracy of the classification systems generated by ID3, C4.5, CBA, Harmony and GARC on the twelve benchmark data sets. The best accuracy values obtained for each of data sets is highlighted in bold print. Table 5 shows that GARC outperforms the traditional classification approaches, *i.e.*, ID3 and C4.5 on six data sets and the associative classification approaches on nine data sets.

Table 5. Accuracy comparison of ID3, C4.5, CBA, Harmony and GARC algorithms

Data set	ID3	C4.5	CBA	Harmony	GARC
Monks1	77.0	75.0	**92.0**	83.0	**92.0**
Monks2	64.0	**65.0**	56.0	48.0	56.0
Monks3	94.0	**97.0**	96.3	82.0	96.3
Spect	65.0	64.0	67.0	-	**68.9**
Pima	71.3	72.9	**73.0**	**73.0**	**73.0**
TicTacToe	83.5	**85.6**	63.1	81.0	65.0
Zoo	**98.0**	92.0	82.2	90.0	90.0
Iris	94.0	94.0	95.3	94.7	**95.4**
Wine	84.8	87.0	89.5	63.0	**89.8**
Glass	64.0	69.1	52.0	**81.5**	64.0
Flare	80.1	84.7	**85.0**	83.0	**85.0**
Pageblocks	92.3	**92.4**	89.0	91.0	89.8

Statistics depicted by Table 5 confirm the fruitful impact of the use of the generic rules. The main reason for this is that Garc classifier contains generic rules with small premises. In fact, this kind of rule allows to classify more objects than those with large premises.

6 Conclusion

In this paper, we introduced a new classification approach called Garc that aims to prune the set of classification rules without jeopardizing the accuracy and even ameliorates the predictive power. To this end, Garc uses generic bases of association rules to drastically reduce the number of associative classification rules. Moreover, it proposes a new score to ameliorate the rules selection for unseen objects. Carried out experiments outlined that Garc is highly competitive in terms of accuracy in comparison with popular classification methods. In the near future, we will investigate new metrics for the rule selection and we will apply Garc approach to a wide range of applications like text categorization and biological applications.

Acknowledgements. We are deeply grateful to Frans Coenen at the university of Liverpool for providing us the discretized UCI data sets and addressing our questions. We also thank Jianyong Wang for providing us the Harmony executable code.

References

1. Zaiane, O., Antonie, M.: On pruning and tuning rules for associative classifiers. In: Ninth International Conference on Knowledge Based Intelligence Information And Engineering Systems (KES'05), Melbourne, Australia (2005) 966–973
2. Xiaoxin Yin, J.H.: CPAR: Classification based on Predictive Association Rules. In: Proceedings of the SDM, San Francisco, CA (2003) 369–376
3. Quinlan, J.R.: C4.5 : Programs for Machine Learning. (1993)
4. Liu, B., Hsu, W., Ma, Y.: Integrating classification and association rule mining. In: Knowledge Discovery and Data Mining. (1998) 80–86
5. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules. In Bocca, J.B., Jarke, M., Zaniolo, C., eds.: Proceedings of the 20th Intl. Conference on Very Large Databases, Santiago, Chile. (1994) 478–499
6. Li, W., Han, J., Pei, J.: CMAR: Accurate and efficient classification based on multiple class-association rules. In: Proceedings of IEEE International Conference on Data Mining (ICDM'01), San Jose, CA, IEEE Computer Society (2001) 369–376
7. Antonie, M., Zaiane, O.: Text Document Categorization by Term Association. In: Proc. of the IEEE International Conference on Data Mining (ICDM'2002), Maebashi City, Japan (2002) 19–26
8. Antonie, M., Zaiane, O.: Classifying Text Documents by Associating Terms with Text Categories . In: Proc. of the Thirteenth Austral-Asian Database Conference (ADC'02), Melbourne, Australia (2002)
9. Quinlan, J., Cameron-Jones, R.: FOIL: A midterm report. In: Proceedings of European Conference on Machine Learning, Vienna, Austria. (1993) 3–20

10. Wang, J., Karypis, G.: HARMONY: Efficiently mining the best rules for classification. In: Proceedings of the International Conference of Data Mining (SDM'05). (2005)
11. Bastide, Y.: Data mining : algorithmes par niveau, techniques d'implantation et applications. Phd thesis, Ecole Doctorale Sciences pour l'Ingénieur de Clermont-Ferrand, Université Blaise Pascal, France (2000)
12. Bastide, Y., Pasquier, N., Taouil, R., Lakhal, L., Stumme, G.: Mining minimal non-redundant association rules using frequent closed itemsets. In: Proceedings of the International Conference DOOD'2000, LNAI, volume 1861, Springer-Verlag, London, UK. (2000) 972–986
13. Gasmi, G., BenYahia, S., Nguifo, E.M., Slimani, Y.: $\mathcal{IGB}$: A new informative generic base of association rules. In: Proceedings of the Intl. Ninth Pacific-Asia Conference on Knowledge Data Discovery (PAKDD'05), LNAI 3518, Hanoi, Vietnam, Springler-Verlag (2005) 81–90
14. Ganter, B., Wille, R.: Formal Concept Analysis. Springer-Verlag (1999)
15. Pasquier, N., Bastide, Y., Taouil, R., Lakhal, L.: Efficient Mining of Association Rules Using Closed Itemset Lattices. Journal of Information Systems **24** (1999) 25–46
16. Kryszkiewicz, M.: Concise representations of association rules. In: Proceedings of Exploratory Workshop on Pattern Detection and Discovery in Data Mining (ESF), 2002, LNAI, volume 2447, Springer-Verlag, London, UK. (2002) 92–109
17. Kryszkiewicz, M.: Representative association rules and minimum condition maximum consequence association rules. In: Proceedings of the Second European Symposium on Principles of Data Mining and Knowledge Discovery (PKDD), 1998, LNCS, volume 1510, Springer-Verlag, Nantes, France. (1998) 361–369
18. Han, J., Kamber., M.: Data Mining : Concepts and Techniques. Morgan Kaufmann. (2001)
19. Hamrouni, T., BenYahia, S., Slimani, Y.: Prince : An algorithm for generating rule bases without closure computations. In Tjoa, A.M., Trujillo, J., eds.: Proceedings of 7th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2005), Springer-Verlag, LNCS 3589, Copenhagen, Denmark. (2005) 346–355